Statistical Inference in the Presence of Ranking Error in Ranked-set Sampling

Omer Ozturk

The Ohio State University

Department of Mathematics
June 22, 2011

Outline

Outline

Outline

1. Ranked Set Sampling
2. Motivation
3. Ranking Model
4. Performance of Estimator
5. Nonparametric Estimate of Judgment Ranking Model
6. Applications
7. Example
8. Conclusion

Outline

Outline

Outline

Outline

Outline

- Select k units at random from a specified population.
- Rank these k units with some expert judgment without measuring them.
- Retain the smallest judged unit and return the others.
- Select the second k units and retain the second smallest unit judged.
- Continue to the process until k ordered units are measured.

Note: These k ordered observations $X_{[1]i}, ..., X_{[k]i}$ are called a cycle.
Note: Process repeated $i = 1, \cdots, n$ cycle to get nk observations.
These *nk* observations are called a standard ranked set sample.

Diagram

Let k=4 and n=3

Cycle	Judgment Rank 1	2	3	4
1	$X_{[1]1}$	.	.	.
	.	$X_{[2]1}$	.	.
	.	.	$X_{[3]1}$	.
	.	.	.	$X_{[4]1}$
2	$X_{[1]2}$	.	.	.
	.	$X_{[2]2}$	.	.
	.	.	$X_{[3]2}$	.
	.	.	.	$X_{[4]2}$

$X_{[1]1}, \cdots, X_{[4]3}$ is called a ranked set sample.

- For each fully measured unit, we need $k - 1$ additional units for ranking.
- Measured units are all independent.
- Under a stable ranking condition, observations from the same judgment class are identically distributed.

Motivating Examples: One-Sample Problem

- Let $X_{[i]j}$, $j = 1, \cdots, n_i$, $i = 1, \cdots, k$ be a ranked set sample from distribution F with unique median θ.
- We wish to test $H_0 : \theta = 0$ against $H_A : \theta \neq 0$.
- A natural nonparametric test statistic for this test is the sing statistic, $S^+ = \sum_{i=1}^{k} \sum_{j=1}^{n_i} I(X_{[i]j} > 0)$.
- Exact null distributing of S^+ involves the convolution of binomial random variables with success probabilities $F_{[i]}(0)$.
- Under perfect ranking $F_{[i]}(0)$, $i = 1, \cdots, k$ are equal to incomplete beta function and the exact null distribution of S^+ can be computed.
- Under imperfect ranking, the size of the test is inflated and the coverage probability of the median confidence interval is deflated.

Estimated type I error rates of the signs test and coverage probabilities of the median confidence interval

k	n	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1.00$
2	5	0.080(0.915)	0.070(0.929)	0.051(0.952)
	7	0.079(0.921)	0.075(0.926)	0.050 (0.954)
3	5	0.112(0.888)	0.091(0.909)	0.047(0.952)
	7	0.102(0.876)	0.084(0.919)	0.053(0.952)

Motivating Examples: Two-sample problem

- Suppose that we wish to test the location shift between $F(y)$ and $G(y) = F(y - \Delta)$. $H_0 : \Delta = \theta_x - \theta_y = 0$ against $H_A : \Delta \neq 0$.
- Assume that we have two data sets, one based on a SRS and the other based on a RSS.
- Let $\bar{T}_{SRS}$ and $\bar{T}_{BW}$ be the rank-sum statistics based on SRS and RSS data.
- The statistic $\bar{T}_{SRS}$ is distribution-free under null hypothesis and its null variance is $1/(12\lambda(1-\lambda))$.
- The limiting null distribution of $\bar{T}_{BW}$ is normal with mean zero and variance

$$
\begin{aligned}
\sigma^2_{BW} &= \xi_{1,0}/\lambda + \xi_{0,1}/(1-\lambda) \\
\xi_{0,1} &= 1/3 - \frac{1}{k}\sum_{i=1}^{k}\left\{\int F_{[i]}(y)dF(y)\right\}^2 \\
\xi_{1,0} &= 1/3 - \frac{1}{q}\sum_{i=1}^{q}\left\{\int F_{[i]}(y)dF(y)\right\}^2.
\end{aligned}
$$

Two-sample problem– Continued

- Under perfect ranking, $\xi_{0,1} = 1/3 - (2k+1)/(6(k+1))$ and $\xi_{1,0} = 1/3 - (2q+1)/(6(q+1))$.
- Under perfect ranking, $N = M$ $k = q$, T_{RSS} is more efficient than T_{SRS}

  $$eff(T_{SRS}, T_{RSS}) = \frac{6(k+1)}{12}.$$

  The efficiency is 1,1.5,2.0,2.5,3.0,3.5 when $k = 1, 2, 3, 4, 5, 6$.
- Under imperfect ranking size of the test is inflated.
- Coverage probabilities of the confidence interval of the shift parameter are deflated.

Type I error rates and coverage probabilities

k	n	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1.00$
2	4	0.116(0.894)	0.091(0.917)	0.066(0.947)
	6	0.106(0.887)	0.078(0.915)	0.052(0.944)
	8	0.111(0.903)	0.084(0.920)	0.061(0.942)
	10	0.111(0.894)	0.079(0.917)	0.053(0.950)
	12	0.103(0.900)	0.081(0.922)	0.051(0.946)
3	4	0.139(0.858)	0.115(0.891)	0.065(0.949)
	6	0.139(0.847)	0.107(0.882)	0.054(0.947)
	8	0.151(0.856)	0.115(0.896)	0.050(0.951)
	10	0.145(0.860)	0.115(0.890)	0.051(0.951)
	12	0.142(0.852)	0.107(0.898)	0.053(0.947)
5	4	0.229(0.780)	0.163(0.829)	0.052(0.944)
	6	0.221(0.776)	0.167(0.833)	0.059(0.947)
	8	0.211(0.781)	0.172(0.838)	0.053(0.949)
	10	0.216(0.776)	0.161(0.840)	0.050(0.944)
	12	0.225(0.787)	0.159(0.848)	0.056(0.952)

Question:

What should we do when we have ranking error?

- Use balanced ranked set sampling if possible.
- Use robust procedures against imperfect ranking.
- Use a ranking model to explain ranking mechanism.
- Estimate this ranking model and use it to calibrate the tests and confidence intervals.

Bohn-Wolfe Model (Bohn and Wolfe,1994 and Frey, 2006)

Bohn-Wolfe Model (Bohn and Wolfe,1994 and Frey, 2006)

$$F_{[i]}(y) = \sum_{s=1}^{k} p_{s,i} F_{(s)}(y), i = 1, \cdots, k,$$

where $0 \leq p_{s,i} \leq 1$ is the probability that the s-th smallest unit is assigned a judgment rank i.

- The matrix $(p)_{s,i}$ defines a stochastic matrix.
- The quality of judgment ranking is controlled by $p_{s,i}$.
- Identity matrix defines a perfect ranking.
- The entries $p_{s,i} = 1/k, s, i = 1, \cdots, k$ define a random ranking.
- Estimate $p_{s,i}$ and use it to draw inference.

Dell-Clutter Model (Dell and Clutter, 1972, David and Levine, 1972)

Dell-Clutter Model (Dell and Clutter, 1972, David and Levine, 1972)

- The true value of a unit, X_i, in a set is modeled through its perceived value $U_i = X_i + \epsilon_i$, where ϵ_i are iid draws from a suitably chosen distribution.
- The sets of (X_i, U_i), $i = 1, \cdots, k$, are ranked with respect to the second component $(X_{[i]}, U_{(i)})$ and the first components are taken as judgment ranked order statistics.
- The quality of ranking is controlled by the noise variable ϵ.
- Monotone Likelihood Ratio model (Fligner and MacEachern, 2006):
  Ranking is performed based on monotone likelihood ratio principal.

Estimation of Bohn-Wolfe Model

- Let $X_{[i]j}$, $i = 1, \cdots, k; j = 1, \cdots, n$ be a ranked set sample from a distribution F and $\hat{F}_{[i]}(y)$ be the empirical cdf of $F_{[i]}$.
- Let $z_{(1)} < ... < z_{(N)}$ be the ordered values of $X_{[i]j}$. Define $u_{[i]j} = \hat{F}_{[i]}(z_j)$, $j = 1, \cdots, N$.
- The distance between the Bohn-Wolfe model and the data can be measured

$$D(\boldsymbol{P}) = \sum_{j=1}^{N} \sum_{i=1}^{k} (u_{[i]j} - \sum_{s=1}^{k} p_{s,i} B_{s,k+1-s}(u_{[i]j}))^2, \tag{1}$$

  where $B_{a,b}(u)$ is the incomplete beta function.
- The number of unknown parameters in $\boldsymbol{P}$ is $M = (k-1)k/2$. Let $\boldsymbol{P}^*$ denote these unknown parameters.

- With appropriate notation, the distance function 1 can be written as

$$D(\boldsymbol{P}^*) = ||\boldsymbol{Y} - \boldsymbol{C}\boldsymbol{P}^*||^2,$$

where $\boldsymbol{Y}$ and $\boldsymbol{C}$ are known kN dimensional vector and $kM \times M$ dimensional matrix, respectively.

- Estimator: We use the minimizer of $D(\boldsymbol{P}^*)$ to estimate the Bohn-Wolfe model

$$\hat{\boldsymbol{P}}^* = (\boldsymbol{C}^\top \boldsymbol{C})^{-1} C^\top \boldsymbol{Y}.$$

- This estimatorv, with direct minimization, may not produce a stochastic matrix. Hence, the estimated ranking model may not be a valid probability model.
- We use constraint quadratic minimization.

Constraint Quadratic Minimization

- We minimize $D(\boldsymbol{P}^*)$ with repect to $\boldsymbol{P}^*$ subject to constraint:
  - Constraint 1: $p_{i,j} \geq 0$ for $i, j = 1, \cdots, k$.
  - Constraint 2: $p_{i,j} \leq 1$ for $i, j = 1, \cdots, k$.
  - Constraint 3: sum in each row (and also column) must be equal to one, $\sum_{j=1}^{k} p_{i,j} = 1$ and $\sum_{i=1}^{k} p_{i,j} = 1$.
- For this minimization, we can use *solve.QP function in R-library quadprog.*

Performance of the Estimator

Consider three different models:

- For $k = 2$

$$P_1 = \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0.90 & 0.1 \\ 0.1 & 0.90 \end{pmatrix},$$

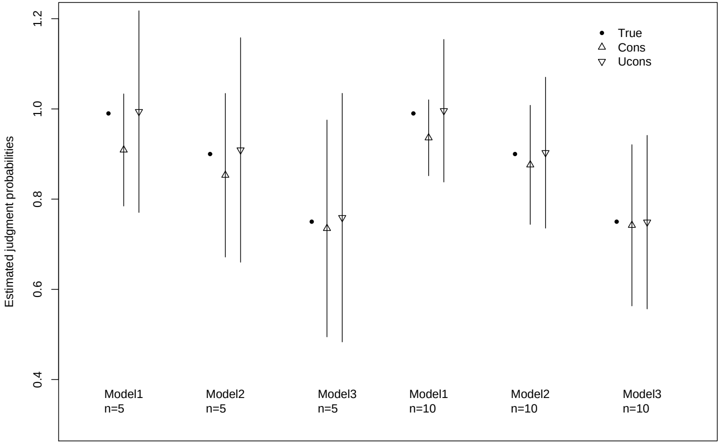
$$P_3 = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}.$$

- For $k = 2$

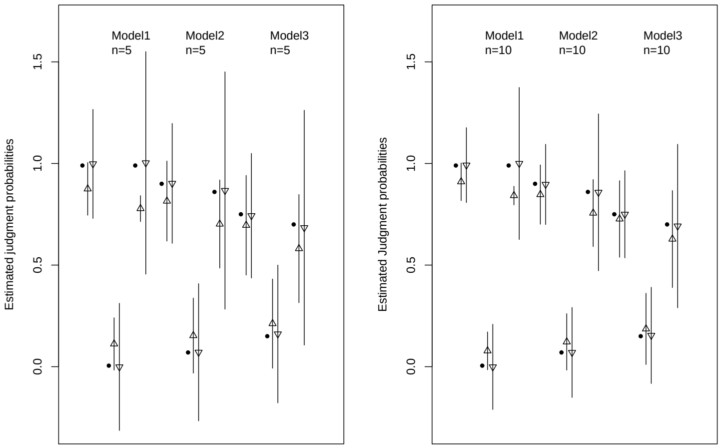
$$P_1 = \begin{pmatrix} .99 & .005 & .005 \\ .005 & .99 & .005 \\ .005 & .005 & .99 \end{pmatrix}, P_2 = \begin{pmatrix} .90 & .07 & .03 \\ .07 & .86 & .07 \\ .03 & .07 & .90 \end{pmatrix},$$

$$P_3 = \begin{pmatrix} 0.75 & 0.15 & 0.10 \\ 0.15 & 0.70 & 0.15 \\ 0.10 & 0.15 & 0.75 \end{pmatrix}.$$

The estimate of judgment ranking probability $p_{1,1}$, $k = 2$



The estimate of judgment ranking probabilities, $p_{1,1}, p_{1,2}, p_{2,2}$, $k = 3$

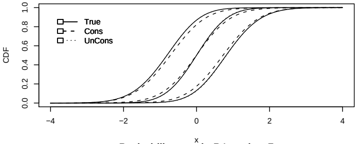


Estimation of judgment class CDF $F_{[i]}$

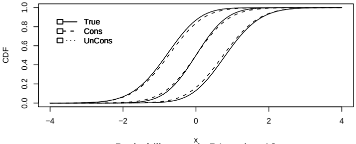
- Assume that the cdf of the underlying distribution F is known.
- From Bohn-wolfe model Judgment class CDF can be estiamted by

$$\hat{F}_{[i]}(y) = \sum_{j=1}^{k} \hat{p}_{i,j} F_{(i)}(y), \quad i = 1, \cdots, k.$$

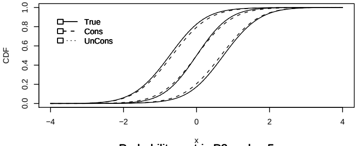
- $\hat{p}_{i,j}$ are the estimates from Bohn-wolfe model.



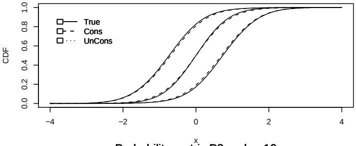
Probability matrix P1 and n=5



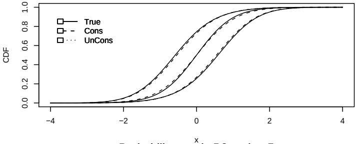
Probability matrix P1 and n=10



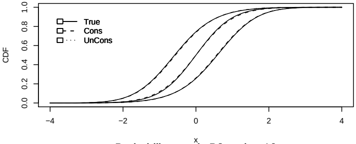
Probability matrix P2 and n=5



Probability matrix P2 and n=10



Probability matrix P3 and n=5



Probability matrix P3 and n=10

Nonparametric Estimate of $F_{[i]}$ (Ozturk, 2007)

- Let $F_{[i]}(y)$, $i = 1, \cdots, k$, be judgment class distributions in a ranked-set sample.
- Any sensible ranking model induces a stochastic ordering between judgment class distributions

$$F_{[1]}(y) \overset{st}{\leq} F_{[2]}(y) \overset{st}{\leq} \cdots \overset{st}{\leq} F_{[k]}(y).$$

- Let $X_{[i]j}$, $i = 1 \cdots, k$, $j = 1, \cdots, n_i$, be a ranked set sample from a distribution F and and let $\hat{F}_{[i]}(y)$ be the empirical cdf of $F_{[i]}(y)$,

$$\hat{F}_{[i]}(y) = \frac{1}{n_i} \sum_{j=1}^{n_i} I(X_{[i]j} \leq y),$$

  where $I()$ is the indicator function.
- Empirical cdf estimator of $F_{[i]}(y)$ does not satisfy this stochastic order restriction.
- **Question:** How do we estimate $F_{[i]}(y)$, $i = 1, \cdots, k$?

Stochastic Ordering Estimator

- Let $z_{(1)} \leq z_{(2)} \leq \cdots \leq z_{(N)}$ be the ordered values of $X_{[i]j}, j = 1, \cdots, n_i; i = 1, \cdots, k$, $\phi_{[i]j} = F_{[i]}(z_{(j)})$ and $\phi = (\phi_{[1]1}, \cdots, \phi_{[k]N})$.
- To find an estimator we want to minimize

$$D(\phi) = \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{n_i} \{\phi_{[i]j} - \hat{F}_{[i]}(z_{(j)}\}^2$$

subject to constraints

$$\phi_{[i]1} \leq \phi_{[i]2} \leq \cdots \leq \phi_{[i]N} \quad (2)$$

$$\phi_{[1]j} \geq \phi_{[2]j} \geq \cdots \geq \phi_{[k]j}. \quad (3)$$

- Let C be the set of real numbers that satisfy the inequalities (2) and (3). Then our estimator is defined as

$$D(\hat{\phi}) = \min_{\phi \in C} D(\phi). \quad (4)$$

- There exist a minimizing set $\hat{\phi} \in C$.

Estimator–Continued

- The estimator can be written in a closed form

$$F^*_{[i]}(z_{(j)}) = \hat{\phi}_{[i]j} = \min_{1 \le r \le i} \max_{i \le s \le k} B_{rs}(j),$$

where $B_{rs}(j) = \frac{\sum_{u=r}^{s} n_u \hat{F}_{[u]}(z_{(j)})}{\sum_{u=r}^{s} n_u}$.

- Under some regularity conditions, $F^*_{[i]}(t)$ uniformly converges almost surely to $F_{[i]}(t)$, $i = 1, \cdots, k$.

Performance of the Estimator

- The proposed estimator satisfies the stochastic order restriction.
- Both empirical cdf and the new estimators are uniformly consistent for $F_{[i]}(t)$. Thus, we expect that both estimators satisfy the stochastic order restriction for large sample sizes.
- For small sample sizes, empirical cdf estimator may violate the stochastic order restriction. In this case, the proposed estimator combines the data across judgment classes where the violation occurs, so we expect it to perform better.
- For small sample sizes, the new estimator has smaller IMSE (Integrated mean square error).
- The proposed estimator has smaller MSE than the empirical cdf estimators.

Efficiency of the estimators, $R_1 = I(E)/I(N)$, $R_2 = I(E)/I(BW)$, $R_3 = I(N)/I(BW)$.

			$\rho = 1$			$\rho = .075$		
k	n	i	R_1	R_2	R_3	R_1	R_2	R_3
2	5	1	1.038	8.960	8.368	1.099	6.173	5.616
	7	1	1.023	9.329	9.115	1.079	5.975	5.539
	10	1	1.011	9.238	9.134	1.058	5.444	5.146
3	5	1	1.057	7.025	6.645	1.169	4.479	3.833
		2	1.128	5.398	4.784	1.337	4.263	3.187
	7	1	1.035	6.872	6.642	1.126	4.346	3.859
		2	1.084	5.289	4.879	1.270	3.853	3.033
	10	1	1.023	6.391	6.244	1.095	3.873	2.827
		2	1.051	5.149	4.897	1.198	3.386	2.827
5	5	1	1.070	4.897	4.578	1.229	2.356	1.917
		2	1.209	2.719	2.249	1.563	2.349	1.502
		3	1.257	2.493	1.983	1.700	2.438	1.434
	10	1	1.034	4.429	4.282	1.143	2.242	1.961
		2	1.103	2.320	2.104	1.401	2.087	1.489
		3	1.133	2.119	1.870	1.490	2.539	1.703

Two-Sample Problem

- Center $X_{[i]j}$ and $Y_{[i]j}$ by subtracting their median. Let $Z_{[i]j}, j = 1, \cdots, n_i + m_i$ be the combined centered X- and Y-observations from the i-judgment class distributions, $i = 1, \cdots, k$.
- We estimate $\xi_{0,1}$ with

$$\hat{\xi}^{N}_{0,1} = \frac{1}{3} - \frac{1}{k}\sum_{i=1}^{k}\left\{\sum_{j=1}^{N}\hat{F}_{[i]}(z_{(j)})d\hat{F}(z_{(j)})\right\}^2$$

$$\hat{\xi}^{BW}_{0,1} = 1/3 - \frac{1}{k}\sum_{i=1}^{k}\left\{\sum_{s=1}^{k}\hat{p}_{s,i}\frac{k+1-i}{k+1}\right\}^2$$

- We estimate $\xi_{1,0}$ and σ_{RSS} in a similar fashion.
- We use the test statistics $\hat{T}_{BW} = \frac{\sqrt{N+M}\{\bar{T}-1/2\}}{\hat{\sigma}_{BW}}$, $\hat{T}_N = \frac{\sqrt{N+M}\{\bar{T}-1/2\}}{\hat{\sigma}_N}$
- We expect that $\hat{T}_{BW}$ and $\hat{T}_N$ have an approximate t-distribution

Estimated Type I error rates. B1- truncated Bown-Wolfe model, B2- Un-truncated BW model, N-nonparametric, P-perfect ranking.

		$\rho = 0.75$				$\rho = 1.00$			
k	n	*B1*	*B2*	*N*	*P*	*B1*	*B2*	*N*	*P*
2	4	.036	.054	.052	.096	.024	.061	.067	.062
	6	.044	.051	.050	.080	.031	.057	.053	.051
	10	.047	.049	.054	.079	.034	.047	.062	.047
	12	.045	.047	.052	.081	.042	.054	.049	.055
3	4	.051	.061	.062	.117	.033	.066	.067	.057
	6	.053	.055	.054	.115	.036	.061	.059	.052
	10	.052	.052	.057	.112	.037	.054	.053	.051
	12	.043	.045	.051	.101	.040	.057	.056	.052
5	4	.071	.069	.062	.167	.033	.067	.067	.053
	6	.059	.056	.056	.169	.034	.060	.066	.054
	10	.050	.053	.058	.159	.038	.058	.056	.055
	12	.051	.053	.050	.153	.033	.052	.050	.049

Estimated coverage probabilities. B1- truncated Bown-Wolfe model, B2- Un-truncated BW model, P-perfect ranking.

		$\rho = 0.75$			$\rho = 1.00$		
k	n	$B1$	$B2$	P	$B1$	$B2$	P
2	4	96.2	94.6	91.7	97.2	93.4	94.7
	6	95.8	95.0	91.5	96.9	94.2	94.4
	10	95.3	95.1	91.7	96.6	95.3	95.0
	12	95.4	95.2	92.2	95.9	94.6	94.6
3	4	95.0	93.9	89.0	96.7	93.5	94.9
	6	94.7	94.5	88.2	96.4	94.1	94.7
	10	94.8	94.8	89.0	96.4	94.6	95.1
	12	95.7	95.4	89.8	96.0	94.3	94.7
5	4	92.9	93.0	82.9	96.9	93.3	94.4
	6	94.7	94.5	88.2	96.6	94.1	94.7
	10	95.0	94.7	84.0	96.2	94.2	94.4
	12	94.9	94.7	84.8	96.6	94.8	95.2

One-Sample Problem

- Let $X_{[i]j}$, $j = 1, \cdots, n_i$, $i = 1, \cdots, k$ be a ranked set sample from distribution F with unique median θ.
- We wish to test $H_0 : \theta = 0$ against $H_A : \theta \neq 0$.
- A natural nonparametric test statistic for this test is the sign statistic, $S^+ = \sum_{i=1}^{k} \sum_{j=1}^{n_i} I(X_{[i]j} > 0)$.
- Exact null distributing of S^+ involves the convolution of binomial random variables with success probabilities $F_{[i]}(0)$.
- Under perfect ranking $F_{[i]}(0)$, $i = 1, \cdots, k$ are equal to incomplete beta function and the exact null distribution of S^+ can be computed.
- Under imperfect ranking, $F_{[i]}(0)$, $i = 1, \cdots, k$ need to be estimated with $\hat{F}_{[i]}(0)$, $i = 1, \cdots, k$.
- With estimated $F_{[i]}(0)$, $i = 1, \cdots, k$, the size of the sing test is stable under imperfect ranking.

Estimated type I error rate of the sing test

k	n	Est.	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1.00$
2	5	BW1	0.058	0.055	0.042
		BW2	0.068	0.066	0.057
		P	0.082	0.070	0.051
2	7	BW1	0.056	0.057	0.045
		BW2	0.060	0.068	0.058
		P	0.079	0.075	0.050
3	5	BW1	0.066	0.064	0.040
		BW2	0.071	0.073	0.060
		P	0.112	0.091	0.047
3	7	BW1	0.057	0.058	0.045
		BW2	0.060	0.062	0.063
		P	0.102	0.084	0.053

Estimated coverage probabilities of the median confidence interval

k	n	Est.	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1.00$
2	5	BW1	94.0	94.6	95.9
		BW2	93.3	93.9	94.6
		P	91.5	92.9	95.2
2	7	BW1	94.4	94.2	95.9
		BW2	94.0	93.3	94.7
		P	92.1	92.6	95.4
3	5	BW1	93.8	94.1	96.4
		BW2	93.4	93.3	95.0
		P	88.8	90.9	95.2
3	7	BW1	94.6	94.7	96.2
		BW2	94.5	93.9	94.2
		P	87.6	91.9	95.2

Example

- An experiment is conducted to test if there is a difference in two spreader settings in in a horticultural experiment.
- The response variable is the percentage area covered by spray deposit on the leaf's upper surface area of apple trees.
- Precise measurement of response is expensive and time consuming with respect to cost of a visual inspection under ultraviolet light.
- Two ranked set samples are collected, one from low spreader settings (A) the other from the high spreader settings (B).
- We wish to test $H_0 : \mu_L - \mu_H = 0$ against $H_A : \mu_L - \mu_H \neq 0$ by using rank-sum statistics.
- The judgment class distributions are estimated under stochastic order restriction.
- The quantity $\xi_{0,1} = \xi_{0,1}$ is estimated $\hat{\xi_{0,1}} = 0.0277$. The test statistic $\hat{T} = 3.635$ and the the P-value is 0.00023. Thus, we reject the null hypothesis.

Rank	Low-volume				
Cycle	1	2	3	4	5
1	0.3	2.8	24.4	5.7	14.3
2	3.9	11.9	12.6	10.5	56.5
3	3.4	11.8	13.0	21.8	29.6
4	5.1	10.4	19.3	21.0	15.0
5	3.2	14.1	13.0	25.0	22.9
6	6.9	7.0	26.0	22.5	28.5
7	10.0	9.1	24.4	13.0	34.7
8	1.2	9.6	6.9	37.3	13.3
9	4.6	6.0	12.6	22.3	27.3
10	2.8	8.3	10.8	21.2	26.1

- We introduced two estimators for the quality of ranking information in a ranked set sample.
- These estimators are used to reduce the effect of imperfect ranking in one and two sample nonparametric statistical procedures.
- These estimated models can easily be extended to other statistical procedures to calibrate the impact of imperfect ranking.